**Estatística descritiva para dados não agrupados**

| | | |
|---|---|---|
| $\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ | $s^2 = \frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{n-1} = \frac{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}{n-1}$ | $s = \sqrt{s^2}$ |
| $c_v = \frac{s}{\overline{x}} \times 100\%$ | $m_k = \frac{\sum_{i=1}^{n}(x_i-\overline{x})^k}{n-1}$ | $c_a = \frac{m_3}{s^3}$ |
| $c_c = \frac{m_4}{s^4}$ | $IQ = Q_{\frac{3}{4}} - Q_{\frac{1}{4}}$ | $k = \frac{IQ}{2\left(Q_{\frac{9}{10}} - Q_{\frac{1}{10}}\right)}$ |
| $G = \frac{\overline{x}-Mo}{s}$ | $G_1 = \frac{3(\overline{x}-Me)}{s}$ | $G_2 = \frac{Q_{\frac{3}{4}} + Q_{\frac{1}{4}} - 2Me}{IQ}$ |

$$Q_p = (1-k)\,x_{(i)} + kx_{(i+1)}, \text{ onde}$$

$$i = \lfloor np+1-p \rfloor, \qquad k = np+1-p-i, \qquad 0 < p < 1,$$

$x_{(i)}$ é o $i$-ésimo valor observado na amostra ordenada e $\lfloor y \rfloor$ é a parte inteira de $y$

Outliers moderados:

$$Q_{\frac{1}{4}} - 3IQ \leqslant x_i < Q_{\frac{1}{4}} - 1,5IQ \quad \text{ou} \quad Q_{\frac{3}{4}} + 1,5IQ < x_i \leqslant Q_{\frac{3}{4}} + 3IQ$$

Outliers severos ou extremos:

$$x_i < Q_{\frac{1}{4}} - 3IQ \quad \text{ou} \quad x_i > Q_{\frac{3}{4}} + 3IQ$$

**Estatística descritiva para dados agrupados**

| | |
|---|---|
| $\overline{x} = \frac{\sum_{i=1}^{k} F_i x_i}{n}$ | $s^2 = \frac{\sum_{i=1}^{k} F_i (x_i-\overline{x})^2}{n-1} = \frac{\sum_{i=1}^{k} F_i x_i^2 - n\overline{x}^2}{n-1}$ |
| Determinação do número de classes: | |
| Fórmula de Sturges: $k = 1 + \lfloor 3,322 \times \log_{10} n \rfloor$, onde $\lfloor y \rfloor$ é a parte inteira de $y$ | É o menor número natural $k$ tal que $2^k \geqslant n$ |

**Técnicas de contagem**

| | |
|---|---|
| $P_n = n!$ | $P'_n = n^n$ |
| ${}^nA_p = \frac{n!}{(n-p)!}$ | ${}^nA'_p = n^p$ |
| ${}^nC_p = \frac{n!}{p!(n-p)!}$ | ${}^nC'_p = \frac{(n+p-1)!}{p!(n-1)!}$ |
| $P(n_1, n_2, \ldots, n_k) = \frac{n!}{n_1! n_2! \cdots n_k!}$ | |

**Teorema de Bayes**

$$\left.\begin{array}{l} A_1 \cup A_2 \cup \cdots \cup A_n = S \\ A_i \cap A_j = \varnothing, \qquad i \neq j \\ P[A_i] > 0, \qquad i = 1, \ldots, n \end{array}\right\} \Rightarrow \begin{array}{l} P[A_j|B] = \frac{P[A_j] \times P[B|A_j]}{P[B]}, \quad j = 1, \ldots, n \\ P[B] = \sum_{i=1}^{n} P[A_i] \times P[B|A_i] \end{array}$$





**Variável aleatória unidimensional**

| Discreta | Contínua |
|---|---|
| $F(x) = P[X \leqslant x] = \sum_{x_j \leqslant x} f(x_j)$ | $F(x) = P[X \leqslant x] = \int_{-\infty}^{x} f(s)\,ds$ |
| $E(X) = \mu = \sum_{i=1}^{n} x_i f(x_i)$ | $E(X) = \mu = \int_{-\infty}^{+\infty} x f(x)\,dx$ |
| $E(X^k) = \sum_{i=1}^{n} x_i^k f(x_i)$ | $E(X^k) = \int_{-\infty}^{+\infty} x^k f(x)\,dx$ |
| $E\left[(X-\mu)^k\right] = \sum_{i=1}^{n} (x_i-\mu)^k f(x_i)$ | $E\left[(X-\mu)^k\right] = \int_{-\infty}^{+\infty} (x-\mu)^k f(x)\,dx$ |
| $Var(X) = \sigma^2 = \sum_{i=1}^{n} (x_i-\mu)^2 f(x_i)$ | $Var(X) = \sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)\,dx$ |
| **Para ambos os casos** | |
| $E(aX \pm b) = aE(X) \pm b$ | $Var(aX \pm b) = a^2 Var(X)$ |
| $E(X \pm Y) = E(X) \pm E(Y)$ | |
| $Var(X \pm Y) = Var(X) + Var(Y)$ se $X$ e $Y$ são independentes | |
| $Var(X) = E(X^2) - [E(X)]^2 = E(X^2) - \mu^2$ | |
| $\sigma = \sqrt{Var(X)}$ | $C_v = \frac{\sigma}{\mu} \times 100\%$ |

**Distribuição Uniforme Discreta**

$$X \sim UniformeDiscreta\{i, \ldots, j\}$$

$$f(x) = \begin{cases} \frac{1}{j-i+1} & , \text{ se } x = i, i+1, \ldots, j \\ 0 & , \text{ se } x \neq i, i+1, \ldots, j \end{cases}$$

$$E[X] = \frac{i+j}{2} \qquad Var[X] = \frac{(j-i+1)^2-1}{12}$$

**Distribuição Bernoulli**

$$X \sim Bernoulli(p)$$

$$f(x) = \begin{cases} p^x (1-p)^{1-x} & , \text{ se } x = 0; 1 \\ 0 & , \text{ se } x \neq 0; 1 \end{cases}$$

$$E[X] = p \qquad Var[X] = p(1-p)$$

**Distribuição Binomial**

$$X \sim Binomial(n; p)$$

$$f(x) = \begin{cases} {}^nC_x p^x (1-p)^{n-x} & , \text{se } \ x = 0, \ldots, n \\ 0 & , \text{se } \ x \neq 0, \ldots, n \end{cases}$$

$$E[X] = np \qquad Var[X] = np(1-p)$$

**Distribuição Multinomial**

$$(X_1, \ldots, X_k) \sim Multinomial(n; p_1; \ldots; p_k)$$

$$P[X_1 = x_1, \ldots, X_k = x_k] = \frac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k},$$

$$x_i = 0, \ldots, n, \qquad i = 1, \ldots, k$$

$$E[X_i] = np_i \qquad Var[X_i] = np_i(1-p_i),\ i = 1, \ldots, k$$

**Distribuição de Poisson**

$$X \sim Poisson(\lambda),\ \lambda > 0$$

$$f(x) = \begin{cases} \frac{\lambda^x \exp\{-\lambda\}}{x!} & , \text{se } \ x = 0, 1, \ldots \\ 0 & , \text{se } \ x \neq 0, 1, \ldots \end{cases}$$

$$E[X] = \lambda \qquad Var[X] = \lambda$$

**Distribuição Binomial Negativa**

$$X \sim BinomialNegativa(r; p)$$

$$f(x) = \begin{cases} {}^{x-1}C_{r-1}\, p^r (1-p)^{x-r} & , \text{se } \ x = r, r+1, \ldots \\ 0 & , \text{se } \ x \neq r, r+1, \ldots \end{cases}$$

$$E[X] = \frac{r}{p} \qquad Var[X] = \frac{r(1-p)}{p^2}$$

| **Distribuição Geométrica** | **Distribuição Hipergeométrica** |
|---|---|
| $X \sim Geométrica\,(p)$ | $X \sim Hipergeométrica\,(N;n;K)$ |
| $f(x)=\begin{cases} p\,(1-p)^{x-1} & ,\text{se } x=1,2,\ldots \\ 0 & ,\text{se } x\neq 1,2,\ldots \end{cases}$ | $f(x)=\begin{cases} \frac{{}^{K}C_{x}\times{}^{N-K}C_{n-x}}{{}^{N}C_{n}} & ,\text{se } x=\max\{0,n-N+K\}, \\ & \ldots,\min\{K,n\} \\ 0 & ,\text{outros valores de } x \end{cases}$ |
| $F(x)=1-(1-p)^{x}$ | |
| $E[X]=\frac{1}{p} \qquad Var[X]=\frac{1-p}{p^2}$ | $E[X]=np \qquad Var[X]=np\,(1-p)\,\frac{N-n}{N-1}$, com $p=\frac{K}{N}$ |
| **Distribuição Uniforme** | **Distribuição Exponencial** |
| $X \sim Uniforme\,(a;b)$ | $X \sim Exponencial\,(\lambda)$ |
| $f(x)=\begin{cases} \frac{1}{b-a} & ,\text{se } a\leqslant x\leqslant b \\ 0 & ,\text{fora do intervalo} \end{cases}$ | $f(x)=\begin{cases} \lambda\exp\{-\lambda x\} & ,\text{se } x>0 \\ 0 & ,\text{se } x\leqslant 0 \end{cases}$ |
| $F(x)=\begin{cases} 0 & ,\text{se } x<a \\ \frac{x-a}{b-a} & ,\text{se } a\leqslant x\leqslant b \\ 1 & ,\text{se } x>b \end{cases}$ | $F(x)=1-\exp\{-\lambda x\},\ x>0$ |
| $E[X]=\frac{a+b}{2} \qquad Var[X]=\frac{(b-a)^2}{12}$ | $E[X]=\frac{1}{\lambda} \qquad Var[X]=\frac{1}{\lambda^2}$ |
| **Distribuição Normal** | **Distribuição Normal Reduzida** |
| $X \sim Normal\,(\mu;\sigma)$ | $X \sim Normal\,(\mu;\sigma) \Rightarrow Z=\frac{X-\mu}{\sigma} \sim Normal\,(0;1)$ |
| $f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}$, | $\varphi(z)=\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}z^2\right\}$, |
| $-\infty<x<+\infty,\ -\infty<\mu<+\infty,\ \sigma>0$ | $-\infty<z<+\infty$ |
| $E[X]=\mu \qquad Var[X]=\sigma^2$ | $E[Z]=0 \qquad Var[Z]=1$ |

**Aditividade da distribuição normal**

$$\left.\begin{array}{c} X_i\,(i=1,\ldots,n)\text{ v. a. independentes} \\ X_i \sim Normal\,(\mu_i;\sigma_i) \end{array}\right\} \Rightarrow S_n=\sum_{i=1}^{n} a_iX_i \sim Normal\left(\sum_{i=1}^{n} a_i\mu_i;\sqrt{\sum_{i=1}^{n} a_i^2\sigma_i^2}\right)$$

$$\left.\begin{array}{c} X_i\,(i=1,\ldots,n)\text{ v. a. independentes} \\ X_i \sim Normal\,(\mu;\sigma) \end{array}\right\} \Rightarrow \begin{array}{c} S_n=\sum_{i=1}^{n} X_i \sim Normal\left(n\mu;\sqrt{n\sigma^2}\right) \\ \overline{X}=\frac{\sum_{i=1}^{n} X_i}{n} \sim Normal\left(\mu;\frac{\sigma}{\sqrt{n}}\right) \end{array}$$

**Teorema Limite Central**

$$\left.\begin{array}{c} X_i\,(i=1,\ldots,n)\ \text{ v. a. i. i. d.} \\ n>30 \\ E[X_i]=\mu \qquad Var[X_i]=\sigma^2 \end{array}\right\} \Rightarrow \begin{array}{c} \frac{\sum_{i=1}^{n} X_i-n\mu}{\sqrt{n\sigma^2}} \mathrel{\dot\sim} Normal\,(0;1) \\ \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \mathrel{\dot\sim} Normal\,(0;1) \end{array}$$





**Variável aleatória bidimensional**

| **Discreta** | **Contínua** |
|---|---|
| $F(x,y)=\sum_{x_i\leqslant x}\sum_{y_j\leqslant y} f(x_i,y_j)=P[X\leqslant x \wedge Y\leqslant y]$ | $F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,dvdu,\ \forall (x,y)\in\mathbb{R}^2$ |
| $f_X(x_i)=\sum_j f(x_i,y_j) \qquad f_Y(y_j)=\sum_i f(x_i,y_j)$ | $f_X(x)=\int_{-\infty}^{+\infty} f(x,y)\,dy \qquad f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)\,dx$ |
| $E[XY]=\sum_i\sum_j x_iy_jf(x_i,y_j)$ | $E[XY]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf(x,y)\,dxdy$ |
| $f_{X\|Y=y_j}(x_i\|y_j)=\frac{f(x_i,y_j)}{f_Y(y_j)}$ com $j$ fixo e $f_Y(y_j)>0$<br>$f_{Y\|X=x_i}(y_j\|x_i)=\frac{f(x_i,y_j)}{f_X(x_i)}$ com $i$ fixo e $f_X(x_i)>0$ | $f_{X\|Y=y}(x\|y)=\frac{f(x,y)}{f_Y(y)}$ com $f_Y(y)>0$<br>$f_{Y\|X=x}(y\|x)=\frac{f(x,y)}{f_X(x)}$ com $f_X(x)>0$ |
| $F_{X\|Y=y_j}(x_i\|y_j)=\sum_{x_i\leqslant x} f_{X\|Y=y_j}(x_i\|y_j)$<br>$F_{Y\|X=x_i}(y_j\|x_i)=\sum_{y_j\leqslant y} f_{Y\|X=x_i}(y_j\|x_i)$ | $F_{X\|Y=y}(x\|y)=\int_{-\infty}^{x}\frac{f(t,y)}{f_Y(y)}dt$<br>$F_{Y\|X=x}(y\|x)=\int_{-\infty}^{y}\frac{f(x,v)}{f_X(x)}dv$ |
| **Para ambos os casos** | |
| $Cov[X,Y]=E[XY]-E[X]\times E[Y]$ | $\rho_{X,Y}=\rho=\frac{Cov[X,Y]}{\sigma_X\sigma_Y}$ |

**Correlação e regressão linear**

| | |
|---|---|
| $s_{xx}=\sum_{i=1}^{n}(x_i-\overline{x})^2=\sum_{i=1}^{n}x_i^2-n\overline{x}^2=(n-1)\,s_x^2$ | $s_{yy}=\sum_{i=1}^{n}(y_i-\overline{y})^2=\sum_{i=1}^{n}y_i^2-n\overline{y}^2=(n-1)\,s_y^2$ |
| $s_{xy}=\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})=\sum_{i=1}^{n}x_iy_i-n\overline{x}\,\overline{y}$ | $cov[x,y]=\frac{s_{xy}}{n-1}$ |
| $r=\frac{cov[x,y]}{s_xs_y}=\frac{s_{xy}}{\sqrt{s_{xx}s_{yy}}}$ | $Y=Y \mid x=\beta_0+\beta_1x+E \qquad \widehat{y}=a+bx$ |
| $b=\frac{s_{xy}}{s_{xx}}=\frac{cov[x,y]}{s_x^2} \qquad a=\overline{y}-b\overline{x}$ | $e_i=y_i-\widehat{y}_i \qquad \frac{e_i-\overline{e}}{s_e}$ |
| $SQR=\sum_{i=1}^{n}(\widehat{y}_i-\overline{y})^2=\frac{s_{xy}^2}{s_{xx}}$ | $SQE=\sum_{i=1}^{n}(y_i-\widehat{y}_i)^2=s_{yy}-bs_{xy}=s_{yy}-\frac{s_{xy}^2}{s_{xx}}$ |
| $SQT=\sum_{i=1}^{n}(y_i-\overline{y})^2=s_{yy}=SQR+SQE$ | $r^2=\frac{SQR}{SQT}=1-\frac{SQE}{SQT}=\frac{s_{xy}^2}{s_{xx}s_{yy}}$ |

| **Função** | **Expressão** | **Transformação** | **Ajustar a recta a** |
|---|---|---|---|
| Exponencial | $\widehat{y}=ab^x$ | $\ln\widehat{y}=\ln a+x\ln b$ | $(x_i,\ln y_i)$ |
| Exponencial | $\widehat{y}=a\exp\{bx\}$ | $\ln\widehat{y}=\ln a+bx$ | $(x_i,\ln y_i)$ |
| Potência | $\widehat{y}=ax^b$ | $\ln\widehat{y}=\ln a+b\ln x$ | $(\ln x_i,\ln y_i)$ |
| Hipérbole | $\widehat{y}=\frac{1}{a+bx}$ | $\frac{1}{\widehat{y}}=a+bx$ | $\left(x_i,\frac{1}{y_i}\right)$ |
| Hipérbole | $\widehat{y}=\frac{x}{a+bx}$ | $\frac{1}{\widehat{y}}=b+a\frac{1}{x}$ | $\left(\frac{1}{x_i},\frac{1}{y_i}\right)$ |
| Inversa | $\widehat{y}=a+\frac{b}{x}$ | $\widehat{y}=a+b\frac{1}{x}$ | $\left(\frac{1}{x_i},y_i\right)$ |
| Logarítmica | $\widehat{y}=a+b\ln x$ | $\widehat{y}=a+b\ln x$ | $(\ln x_i,y_i)$ |
| Curva $S$ | $\widehat{y}=\exp\left\{a+\frac{b}{x}\right\}$ | $\ln\widehat{y}=a+b\frac{1}{x}$ | $\left(\frac{1}{x_i},\ln y_i\right)$ |
| Crescimento | $\widehat{y}=\exp\{a+bx\}$ | $\ln\widehat{y}=a+bx$ | $(x_i,\ln y_i)$ |

## Intervalos de confiança para um parâmetro e dois parâmetros (amostras independentes)

| Parâmetros a estimar | $\sigma^2$ conhecido? | Tipo de população | Dimensão da amostra | Variável fulcral e correspondente distribuição amostral |
|---|---|---|---|---|
| $\mu$ | Sim | Normal | Qualquer | $\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0;1)$ |
| | | Outra | $n>30$ | $\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \dot\sim N(0;1)$ |
| | Não | Qualquer | $n>30$ | $\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}} \dot\sim N(0;1)$ |
| | | Normal | Qualquer | $\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1}$ |
| $p$ | - - - | Bernoulli | $n>30$ | $\frac{\hat{P}-p}{\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}} \dot\sim N(0;1)$ |
| $\sigma^2$ | - - - | Normal | Qualquer | $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ |
| $\mu_1-\mu_2$ | $\sigma_1^2$ e $\sigma_2^2$ conhecidas | Normais | Quaisquer | $\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0;1)$ |
| | | Outras | $n_1>30$ e $n_2>30$ | $\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \dot\sim N(0;1)$ |
| | $\sigma_1^2$ e $\sigma_2^2$ desconhecidas | Quaisquer | $n_1>30$ e $n_2>30$ | $\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} \dot\sim N(0;1)$ |
| | $\sigma_1^2$ e $\sigma_2^2$ desconhecidas $\left(\sigma_1^2=\sigma_2^2\right)$ | Normais | Quaisquer | $\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \sim t_{n_1+n_2-2}$ |
| | $\sigma_1^2$ e $\sigma_2^2$ desconhecidas $\left(\sigma_1^2\neq\sigma_2^2\right)$ | Normais | Quaisquer | $\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} \sim t_r$<br>sendo $r$ o número natural mais próximo de<br>$r^*=\frac{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)^2}{\frac{1}{n_1-1}\left(\frac{S_1^2}{n_1}\right)^2+\frac{1}{n_2-1}\left(\frac{S_2^2}{n_2}\right)^2}$ |
| $p_1-p_2$ | - - - | Bernoulli | $n_1>30$ e $n_2>30$ | $\frac{(\hat{P}_1-\hat{P}_2)-(p_1-p_2)}{\sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n_1}+\frac{\hat{P}_2(1-\hat{P}_2)}{n_2}}} \dot\sim N(0;1)$ |
| $\frac{\sigma_2^2}{\sigma_1^2}$ | - - - | Normais | Quaisquer | $\frac{S_1^2}{S_2^2}\times\frac{\sigma_2^2}{\sigma_1^2} \sim F(n_1-1;n_2-1)$ |





## Testes de hipóteses para um parâmetro e dois parâmetros (amostras independentes)

| Parâmetros a estimar | $\sigma^2$ conhecido? | Tipo de população | Dimensão da amostra | Estatística de teste e correspondente distribuição amostral |
|---|---|---|---|---|
| $\mu$ | Sim | Normal | Qualquer | $Z_0=\frac{\overline{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}} \sim N(0;1)$ |
| | | Outra | $n>30$ | $Z_0=\frac{\overline{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}} \dot\sim N(0;1)$ |
| | Não | Qualquer | $n>30$ | $Z_0=\frac{\overline{X}-\mu_0}{\frac{S}{\sqrt{n}}} \dot\sim N(0;1)$ |
| | | Normal | Qualquer | $T_0=\frac{\overline{X}-\mu_0}{\frac{S}{\sqrt{n}}} \sim t_{n-1}$ |
| $p$ | - - - | Bernoulli | $n>30$ | $Z_0=\frac{\hat{P}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \dot\sim N(0;1)$ |
| $\sigma^2$ | - - - | Normal | Qualquer | $Q_0=\frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}$ |
| $\mu_1-\mu_2$ | $\sigma_1^2$ e $\sigma_2^2$ conhecidas | Normais | Quaisquer | $Z_0=\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)_0}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0;1)$ |
| | | Outras | $n_1>30$ e $n_2>30$ | $Z_0=\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)_0}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \dot\sim N(0;1)$ |
| | $\sigma_1^2$ e $\sigma_2^2$ desconhecidas | Quaisquer | $n_1>30$ e $n_2>30$ | $Z_0=\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)_0}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} \dot\sim N(0;1)$ |
| | $\sigma_1^2$ e $\sigma_2^2$ desconhecidas $\left(\sigma_1^2=\sigma_2^2\right)$ | Normais | Quaisquer | $T_0=\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)_0}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \sim t_{n_1+n_2-2}$ |
| | $\sigma_1^2$ e $\sigma_2^2$ desconhecidas $\left(\sigma_1^2\neq\sigma_2^2\right)$ | Normais | Quaisquer | $T_0=\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)_0}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} \sim t_r$<br>sendo $r$ o número natural mais próximo de<br>$r^*=\frac{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)^2}{\frac{1}{n_1-1}\left(\frac{S_1^2}{n_1}\right)^2+\frac{1}{n_2-1}\left(\frac{S_2^2}{n_2}\right)^2}$ |
| $p_1-p_2$ | - - - | Bernoulli | $n_1>30$ e $n_2>30$ | $Z_0=\frac{(\hat{P}_1-\hat{P}_2)-(p_1-p_2)_0}{\sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n_1}+\frac{\hat{P}_2(1-\hat{P}_2)}{n_2}}} \dot\sim N(0;1)$ |
| $\frac{\sigma_2^2}{\sigma_1^2}$ | - - - | Normais | Quaisquer | $F_0=\frac{S_1^2}{S_2^2}\times\left(\frac{\sigma_2^2}{\sigma_1^2}\right)_0 \sim F(n_1-1;n_2-1)$ |

### Testes de hipóteses para dois parâmetros (amostras emparelhadas)

| Parâmetros a estimar | $\sigma^2$ conhecido? | Tipo de população das diferenças | Dimensão da amostra | Estatística de teste e correspondente distribuição amostral |
|---|---|---|---|---|
| $\mu_D$ | Sim | Normal | Qualquer | $Z_0 = \frac{\overline{D}-(\mu_D)_0}{\frac{\sigma_D}{\sqrt{n}}} \sim N(0;1)$ |
| | | Outra | $n > 30$ | $Z_0 = \frac{\overline{D}-(\mu_D)_0}{\frac{\sigma_D}{\sqrt{n}}} \dot\sim N(0;1)$ |
| | Não | Qualquer | $n > 30$ | $Z_0 = \frac{\overline{D}-(\mu_D)_0}{\frac{S_D}{\sqrt{n}}} \dot\sim N(0;1)$ |
| | | Normal | Qualquer | $T_0 = \frac{\overline{D}-(\mu_D)_0}{\frac{S_D}{\sqrt{n}}} \sim t_{n-1}$ |

### Testes de aderência ou de qualidade de ajuste

Teste de ajustamento de Kolmogorov-Smirnov com correcção de Lilliefors:

A hipótese nula é rejeitada quando $D_{\text{observado}} \geqslant D_{\text{crítico};\alpha}$

Teste de ajustamento de Shapiro-Wilk:

A hipótese nula é rejeitada quando $W_{\text{observado}} \leqslant W_{\text{crítico};\alpha}$

### Testes de homocedasticidade

Teste de Levene:

A hipótese nula é rejeitada quando $F_{\text{observado}} \geqslant F(k-1; n-k; 1-\alpha)$

sendo $k$ o número de amostras e $n = n_1 + n_2 + \cdots + n_k$

Teste de Bartlett:

A hipótese nula é rejeitada quando $Q_{\text{observado}} \geqslant \chi^2_{k-1;1-\alpha}$

sendo $k$ o número de amostras

### p-values

| Teste bilateral | Teste unilateral à direita | Teste unilateral à esquerda |
|---|---|---|
| $2 \times P[Z \geqslant \lvert V.E.T.\rvert]$ | $P[Z \geqslant V.E.T.]$ | $P[Z \leqslant V.E.T.]$ |
| $2 \times P[T \geqslant \lvert V.E.T.\rvert]$ | $P[T \geqslant V.E.T.]$ | $P[T \leqslant V.E.T.]$ |
| $2 \times \min\{P[\chi^2 \leqslant V.E.T.]; P[\chi^2 \geqslant V.E.T.]\}$ | $P[\chi^2 \geqslant V.E.T.]$ | $P[\chi^2 \leqslant V.E.T.]$ |
| $2 \times \min\{P[F \leqslant V.E.T.]; P[F \geqslant V.E.T.]\}$ | $P[F \geqslant V.E.T.]$ | $P[F \leqslant V.E.T.]$ |

$V.E.T.$ - valor da estatística de teste.





### Inferência para o coeficiente de correlação, parâmetros da recta de regressão e previsão

| Parâmetros a estimar | Variável fulcral e correspondente distribuição amostral | Estatística de teste e correspondente distribuição amostral |
|---|---|---|
| $\rho$ | $\frac{Z_R-Z_\rho}{\frac{1}{\sqrt{n-3}}} \dot\sim N(0;1)$ sendo a transformada de Fisher $z_r = \frac{1}{2}\ln\left(\frac{1+r}{1-r}\right)$ e $r = \frac{\exp\{2z\}-1}{\exp\{2z\}+1}$ | $T_0 = R\sqrt{\frac{n-2}{1-R^2}} \sim t_{n-2}$ ($\rho$ próximo de zero e $n \geqslant 3$) |
| | | $Z_0 = \frac{Z_R-(Z_\rho)_0}{\frac{1}{\sqrt{n-3}}} \dot\sim N(0;1)$ ($\rho$ qualquer e $n \geqslant 3$) |
| $\beta_0$ | $\frac{A-\beta_0}{S_{y\cdot x}\sqrt{\frac{1}{n}+\frac{\overline{X}^2}{S_{xx}}}} \sim t_{n-2}$ | $T_0 = \frac{A-(\beta_0)_0}{S_{y\cdot x}\sqrt{\frac{1}{n}+\frac{\overline{X}^2}{S_{xx}}}} \sim t_{n-2}$ |
| $\beta_1$ | $\frac{B-\beta_1}{\frac{S_{y\cdot x}}{\sqrt{S_{xx}}}} \sim t_{n-2}$ | $T_0 = \frac{B-(\beta_1)_0}{\frac{S_{y\cdot x}}{\sqrt{S_{xx}}}} \sim t_{n-2}$ |
| $Y_0$ | $\frac{\widehat{Y}_0-Y_0}{S_{y\cdot x}\sqrt{1+\frac{1}{n}+\frac{(x_0-\overline{X})^2}{S_{xx}}}} \sim t_{n-2}$ | $T_0 = \frac{\widehat{Y}_0-(Y_0)_0}{S_{y\cdot x}\sqrt{1+\frac{1}{n}+\frac{(x_0-\overline{X})^2}{S_{xx}}}} \sim t_{n-2}$ |
| $\mu_{Y_0}$ | $\frac{\widehat{\mu}_{Y_0}-\mu_{Y_0}}{S_{y\cdot x}\sqrt{\frac{1}{n}+\frac{(x_0-\overline{X})^2}{S_{xx}}}} \sim t_{n-2}$ | $T_0 = \frac{\widehat{\mu}_{Y_0}-(\mu_{Y_0})_0}{S_{y\cdot x}\sqrt{\frac{1}{n}+\frac{(x_0-\overline{X})^2}{S_{xx}}}} \sim t_{n-2}$ |

| | | |
|---|---|---|
| $s_{y\cdot x} = \sqrt{\frac{\sum_{i=1}^{n}(y_i-\widehat{y}_i)^2}{n-2}} = \sqrt{\frac{SQE}{n-2}} = \sqrt{\frac{s_{yy}-bs_{xy}}{n-2}}$ | $s_A = s_{y\cdot x}\sqrt{\frac{1}{n}+\frac{\overline{x}^2}{s_{xx}}}$ | $s_B = \frac{s_{y\cdot x}}{\sqrt{s_{xx}}}$ |

### Curvas de calibração em análise instrumental

| Parâmetro a estimar | Variável fulcral e correspondente distribuição amostral |
|---|---|
| $X_0$ | $\frac{\widehat{X}_0-X_0}{\frac{S_{y\cdot x}}{B}\sqrt{\frac{1}{m}+\frac{1}{n}+\frac{(y_0-\overline{Y})^2}{B^2 S_{xx}}}} \sim t_{n-2}$, com $y_0 = \frac{\sum_{i=1}^{m} y_i}{m}$ |

### Análise de variância - ANOVA

A hipótese nula é rejeitada quando $F_{\text{observado}} \geqslant F(k-1; n-k; 1-\alpha)$

sendo $k$ o número de amostras e $n = n_1 + n_2 + \cdots + n_k$

### Tabela ANOVA

| Fonte de variação | Somas de quadrados | Graus de liberdade | Média de quadrados | Estatística de teste $F$ |
|---|---|---|---|---|
| Regressão | $SQR$ | 1 | $MQR = \frac{SQR}{1}$ | $F_0 = \frac{MQR}{MQE} \sim F(1; n-2)$ |
| Erros | $SQE$ | $n-2$ | $MQE = \frac{SQE}{n-2}$ | |
| Total | $SQT$ | $n-1$ | | |

A hipótese nula $\beta_1 = 0$ é rejeitada quando $F_{\text{observado}} \geqslant F(1; n-2; 1-\alpha)$.